IJANR
International Journal of Agriculture and Natural Resources

Int. J. Agric. Nat. Resour.49(1):22-35. 2022
www.ijanr.cl

BIOTECHNOLOGY

DOI 10.7764/ijanr.v49i1.2357

RESEARCH PAPER

Root transcriptome of Allium mongolicum Regel under drought stress conditions

Ning Tang[#], Quan-Zhan Chen[#], Bian-Jiang Zhang*, Ping Yang, and Li-Ke Wang

Nanjing Xiaozhuang University, School of Food Science. 3601 Hongjing Ave, Nanjing, Jiangsu, 211171 China.

Abstract

N. Tang, Q.- Z. Chen, B.-J. Zhang, P. Yang, and L.-K. Wang. 2022. Root transcriptome of Allium mongolicum Regel under drought stress conditions. Int. J. Agric. Nat. Resour. 22-35. *Allium mongolicum Regel* is a desert plant that shows great resistance to wind erosion, drought, and low temperatures and is widely distributed in the desert lands of China. Understanding the molecular mechanisms underlying its drought resistance is essential for uncovering drought resistance genes and improving its beneficial traits. Here, a de novo RNA-Seq assembly analysis was conducted using the roots of 1-month-old seedlings under drought stress. Using pairwise comparisons of untreated plants (CK), 48-hour drought plants (G48h), and 96-hour drought plants (G96h), in total, 2,211 differentially expressed unigenes (DEGs) were obtained. Furthermore, using functional annotation, these DEGs were mainly involved in the plasma membrane, photosynthesis, and secondary carbohydrate metabolism. Moreover, genes involved in the ABA-mediated signaling pathway and secondary metabolism were upregulated in the roots. These results suggest that desert plants may use signaling and secondary metabolic pathways as adaptive responses to drought stress. Collectively, this work could help elucidate the molecular mechanisms underlying the ability of *A. mongolicum Regel* to respond to drought stress and aid in the selection of novel drought tolerance genes.

Keywords: *Allium mongolicum*, drought stress, molecular mechanism, root, transcriptome.

Introduction

Desert regions currently cover at least 35% of the Earth's land surface (Pointing & Belnap, 2012) and are still expanding due to ongoing global warming (Rodriguezcaballero et al., 2018). The desert areas in northern China account for ~17.9% of the total land area in China (Shao et al., 2016), representing a major constraint in agricultural development. Desert plants, such as *Ammopiptan thusmongolicus* (X. Sun et al., 2012), *Zygophyllum xanthoxylum* (Hu et al., 2016), and *A. mongolicum Regel,* are valuable genetic resources in desert areas. Conserving these desert plants is critical for global efforts aiming to curb desertification and prevent the development of fragile ecosystems in arid and semiarid regions as they can lead to further deterioration (Zhou et al., 2012). Moreover, by understanding the drought-response mechanisms

Received Jul 25, 2021. Accepted Dec 15, 2021
# Authors contributed equally to this work.
*Corresponding author: zhangbjiang@njxzc.edu.cn

of desert plants, breeders could be better able to promote drought resistance in other plants.

Plants survive under water-deficit conditions via a series of physiological, cellular, and molecular processes (Shinozaki & Yamaguchi-Shinozaki, 2006). In recent years, given the development of next-generation sequencing (NGS) and high-throughput RNA sequencing, many drought-related genes, enzymes, transcription factors, receptors, lncRNAs, and secondary metabolites have been identified in plants (Qin et al., 2017; Khan et al., 2017; Wang et al., 2010). These stress-response genes tend to encode enzymes regulating osmotic pressure, reactive oxygen species scavengers, and chaperones that protect the integrity of cell membranes and ensure ion transport/balance (Sultana et al., 2016). Additionally, secondary metabolites, such as brassinosteroids, represent an integral component of the stress response to drought.

A. mongolicum Regel (*Amaryllidaceae*), also known as Mongolia leek, wild onion, and mountain onion, is a xeromorphic, bulbous, herbaceous desert plant grown in high-altitude desert steppes and desert areas. *A. mongolicum Regel* is widely distributed in desert lands in Qinghai, Gansu, Xinjiang, and Inner Mongolia and is especially distributed in the western areas of Xilingol, Ordos, and Alxain Inner Mongolia (Muqier et al., 2017). *A. mongolicum Regel* is resistant to wind erosion, drought, and low temperatures and, thus, is important for preventing wind erosion and maintaining and improving regional ecological environments (Zhou et al., 2012).

However, due to the limited genomic information available, studies investigating *A. mongolicum Regel* have mainly focused on its food value and medical functions (Muqier et al., 2017), and genetic studies investigating its agronomically important characteristics, such as drought tolerance, have not been carried out. Furthermore, the drought-response genes and their genetic expression profiles under drought stress in *A. Mongolicum Regel* have rarely been reported. Thus, a deeper understanding of the genetic control of this plant's adaptation to desert environments could be beneficial.

As an economic crop widely distributed in desert areas, *A. mongolicum Regel* has many important agronomic characteristics that many crops do not have, such as drought tolerance and resistance to wind erosion. Despite its economic importance in desert areas, the genomic information of *A. mongolicum Regel* is still limited. To expand the transcript catalog of *A. mongolicum Regel* and identify candidate genes involved in drought tolerance, large-scale transcriptome sequencing of the roots of polyethylene glycol (PEG)-treated *A. mongolicum Regel* was performed using the Illumina HiSeq™ 2000 system and de novo RNA-Seq assembly. The primary aim of this work was to study the expression divergence of drought-induced genes in *A. Mongolicum Regel* while contrasting their phenotypic response to drought. The obtained information provides new insight into the genetic basis of drought tolerance in *A. Mongolicum Regel* and facilitates the identification of potential targets to improve drought tolerance.

Methods

Sample preparation

The sample preparation and drought simulation were performed according to the methods described by Zhou et al. (2012). Generally, seeds of *A. mongolicum Regel* (collected from the desert region in Ningxia Autonomous Region, China) were soaked in water for 12 hours at 25 °C and then sown in petri dishes (9 cm diameter) under a photosynthetic photon flux density of 120 μmol m^{-2} s^{-1} for 30 days with a relative humidity of 35% and a temperature of 25 °C. The plantlets were watered every 3 days with half-strength Hoagland's solution.

One month after germination, the seedlings were divided into three groups. The first group served

as a control (CK) without any treatment; the other two groups were transferred into a nutrient solution containing 20% PEG-6000 (w/v) with the roots totally immersed; one group was cultured for 48 hours (G48h), and the other group was cultured for 96 hours (G96h). After the treatment, the roots of all samples were harvested, and RNA extraction was performed. Each sample-preparation procedure was repeated three times.

RNA isolation, cDNA library construction, and sequencing

The total RNA of nine samples of CK and two treatments was isolated using TRIzol reagent (Life Technologies, Inc., Gaithersburg, MD, USA) according to the manufacturer's protocol. The extracted total RNA was quantified with a Nano Drop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and qualified by agarose gel electrophoresis. The cDNA library construction and sequencing were conducted by Vazyme Biotech Co., Ltd. (Vazyme Biotech, Nanjing, China). In brief, the total RNA was first purified with RNase free DNase, enriched with mRNA Capture Beads (Vazyme Biotech, Nanjing, China), and subsequently fragmented. cDNA was synthesized using random hexamer primers using fragmented mRNA as a template. Then, double-strand cDNA was purified with VAHTSTM DNA Clean Beads (Vazyme Biotech, Nanjing, China) and end-repaired, and an A-tail and sequencing adaptor were added. Proper fragments were selected, polymerase chain reaction (PCR)-amplified and purified to create the final cDNA library. Finally, the library was sequenced using Illumina HiSeq™ 2000 at Vazyme Biotech Co., Ltd. Before assembly, raw reads with adaptors and unidentified nucleotides and low-quality reads were filtered. Then, the clean reads were assembled to obtain unigenes using Trinity software as previously described (Manfred Grabherr et al., 2011a). Reads demonstrating an extreme signal intensity due to the hardware of the sequencing instrument, a low overall quality, and a length <20 and those obtained from ncRNA were removed. We removed bases with a mass Q<10 at the 3' end and vague N bases and connection sequences. The raw sequencing data were deposited in the National Center for Biotechnology Information (NCBI) database under accession number PRJNA526320.

Unigene annotation and classification

For the unigene function annotation, the clean reads were assembled into the NCBI nonredundant protein database (NR), Kyoto Encyclopedia of Genes and Genomes (KEGG), Swiss-Prot, and Clusters of Orthologous Groups (COG) with an E-value threshold of $1.0e^{-5}$. The sequence direction was determined by the best BLAST results. For the unigenes that did not have results consistent with the above databases, ESTScan software was used to determine their sequence direction (Iseli et al., 1999).

For the functional annotation, the unigenes were subjected to a Gene Ontology (GO) analysis using the Blast2go program (Conesa et al., 2005); the GO functional classification was performed using WEGO software (Ye et al., 2006). The pathway assignments were determined based on the KEGG pathway database using BlastX with an E-value threshold of $1.0e^{-5}$.

Analysis of differentially expressed genes (DEGs)

First, the clean reads of all materials were separately aligned to de novo assembled transcriptomes using Bowtie2. The transcript abundance was normalized using the RSEM package, and the fragments per kilobase of transcripts per million fragments mapped (FPKM) values were estimated. The analysis of the differential expression between CK and G48h, CK and G48h, and G48h and G96h was performed separately. The DE analysis was conducted using three different R packages

(DESeq, edgeR, and NOISeq) at values of var=0.05 and $P<0.05$. A cluster analysis of the gene expression patterns was performed using cluster software (De Hoon et al., 2004) and Java Tree view software (Saldanha, 2004). The GO annotation and KEGG pathway analysis of the DEGs were performed as described in the "unigene annotation and classification" section. GO enrichment was tested using Fisher's exact test with a cutoff threshold of FDR<0.05.

Quantitative real-time PCR (qPCR): Validation of differentially expressed genes

Total RNA isolated from Allium seedling roots at different treatment times was first treated with RNase free DNase I (Promega Corporation, Madison, WI, USA) and then purified with phenol: chloroform. The first strand cDNA was synthesized using TransScript® One-Step gDNA Removal and diluted 10 times. qPCR was performed on a CFX Connects real-time system (Bio–Rad, Hercules, CA, USA) using diluted cDNA as a template. Actin was used as an internal control gene; gene-specific primers were designed for qPCR. The relative fold change values were calculated using the ΔΔCt method; each experiment included three biological replicates and two technical replicates. The associated genes are as follows: chlorophyll a-b binding protein (unigene52938), *SWEET14* (unigene46076), lipid transfer protein (LTP) (unigene20794, unigene20795 and unigene20798), LEA (Late Embryogenesis Abundant) protein (unigene39451), MYB family transcription factor (unigene36804, unigene22068, and unigene46936), and actin (unigene2449, unigene74845, unigene74846, unigene74847, unigene74848, and unigene74849).

Results

Illumina sequencing and de Novo assembly

To study the mechanism of drought tolerance in *A. mongolicum Regel*, RNA was extracted from seedlings treated with nutrient solution containing 20% PEG-6000 for 48 hours and 96 hours; samples in nutrient solution without PEG-6000 were used as a control. Using an Illumina HiSeq™ 2000 system, an average of 43,803,484 raw reads in CK, 46,494,204 raw reads in G48h, and 45,694,268 raw reads in G96h were obtained. Following a stringent quality check and data filtering, high-quality clean reads (Q20>96%, Q30>91) with 42% GC content in each sample were obtained (Table 1).

Table 1. Summary of the sequencing reads from the *A. mongolicus* transcriptome under different drought stress conditions.

Samples	Total Raw Reads	Total Clean Reads	Total Clean Nucleotides (nt)	Q20 Percentage	Q30 Percentage	N Percentage	GC Percentage
CK1	36,843,898	35,782,006	5,367,300,900	96.27%	91.39%	0.00%	42.58%
CK2	47,550,616	46,178,500	6,926,775,000	96.73%	92.18%	0.00%	42.77%
CK3	47,015,938	45,598,224	6,839,733,600	96.95%	92.65%	0.00%	43.06%
Average	43,803,484	42,519,577	6,377,936,500	96.65%	92.07%	0.00%	42.8%
G48h1	47,472,846	46,125,034	6,918,755,100	97.28%	93.36%	0.00%	42.95%
G48h2	47,442,188	46,125,034	6,918,755,100	97.17%	93.10%	0.00%	42.92%
G48h3	44,567,580	43,305,152	6,495,772,800	96.97%	92.71%	0.00%	42.76%
Average	46,494,205	45,185,073	6,777,761,000	97.14%	93.06%	0.00%	42.88%
G96h1	45,677,662	44,377,786	6,656,667,900	97.02%	92.83%	0.00%	42.83%
G96h2	47,791,502	46,497,656	6,974,648,400	97.15%	92.92%	0.00%	42.84%
G96h3	43,613,640	42,416,718	6,362,507,700	97.02%	93.17%	0.00%	42.39%
Average	45,694,268	44,430,720	6,664,608,100	97.06%	92.97%	0.00%	42.69%

Note: The Q20 or Q30 percentage represents the ratio of clean reads with lengths longer than 20 or 30, respectively.
The N percentage represents the percentage of reads containing uncertain bases after filtering.
The GC percentage represents the percentage of G+C in total bases.
Total clean nucleotides = Total clean reads1 × Reads1 size + Total clean reads2 × Reads2 size+…+Total clean reads n× Reads n size.

As the full genome sequence of *A. mongolicum Regel* is still not available in public databases, the clean reads were assembled using Trinity software (Manfred G. Grabherr et al., 2011b) de novo into an average total length >60 million contigs (≥200 bp), a mean length >334, and an average N50 >510 in each sample (Table 2). All assembled contigs were >200 bp in size, and the maximum length of a contig was >3,000 bp. Meanwhile, no significant differences in the length distribution of the assembled contigs were discovered between CK and the two treatments (Figure 1A).

After clustering by Chrysalis, in total, 203,443 unigenes with a total length of 182,572,023 bp, a mean length of 897 bp, and an N50 of 1,523 were obtained. The length distribution of the genes is shown in Table 2. Similar to the contig length distribution, there were no significant differences between CK and the treatments in the unigene length (Figure 1B).

Functional annotation and classification

For the annotation, the assembled unigenes were first blasted against six different databases using the BLAST algorithm and a cutoff E-value of 10^{-5} (Table 3). The results indicated that 81,322 of the 85,653 unigenes (94.94%) could be annotated based on the sequences in the nonredundant (NR) database, while in the other five databases, such as the Swiss-Prot, KEGG, Clusters of orthologous Group of proteins (COG), and GO databases, there was a lower similarity compared with the NR database (Table 3).

Based on the statistical analysis results, the distributions of the top hits in the NR database showed that 47.2% and 10.7% of the unigenes demonstrated significant homology (E-value <1E-45) or high similarity (>80%) with the sequences in the NR database (Figures 2A and 2B). Furthermore, 52.8% of the top BLAST hit unigenes were found in the following species: *Vitis vinifera* (18.6%), *Oryza sativa japonica* (8.0%), *Amygdalus persica* (5.8%), *Populus balsamifera* subsp. *Trichocarpa* (5.3%), *Brachypodium distachyon* (5.1%), *Ricinus communis* (5.0%), and *Zea mays* (5.0%) (Figure 2C).

GO Classification and Functional Annotation of all Unigenes

For the annotation of the assembled unigenes, GO and COG analyses were performed. In total, 85,653 annotated unigenes were assigned to one or more GO terms. These unigenes were categorized into three main categories (biological process, cellular component, and molecular function) and 56 subcategories. Meanwhile, the biological process category contained most unigenes. Further

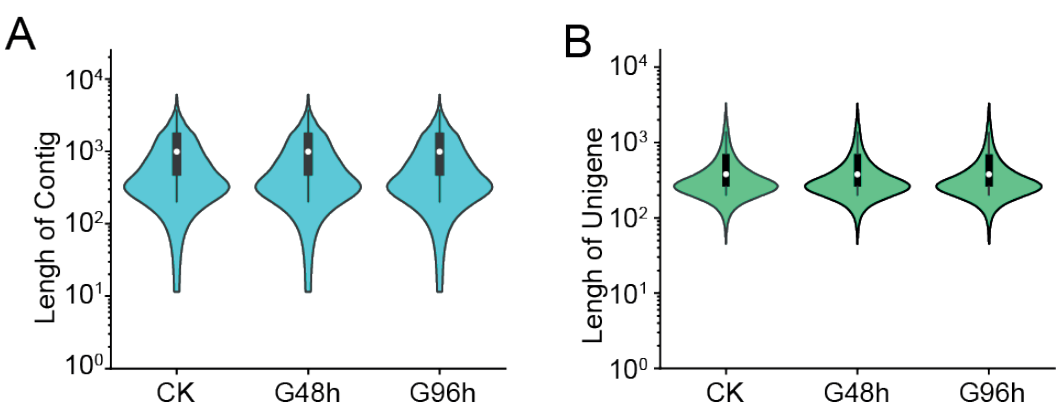

Figure 1. Histograms of the contig and unigene length distribution in the transcriptome of *A. mongolicus* root samples. (A) contigs; (B) unigenes. The X-axis indicates the length range of contigs with sizes ranging from 200 nt to >3,000 nt. CK: control; G48h: treated with 20% PEG-6000 for 48 hours; G96h: treated with 20% PEG-6000 for 96 hours.

Table 2. Summary of the *A. mongolicus* transcriptome.

Sample		Total Number	Total Length (nt)	Mean Length (nt)	N50	Total Consensus Sequences	Distinct Clusters	Distinct Singletons
Contig	CK1	147,266	52,096,679	354	683			
	CK2	196,210	65,455,637	334	562			
	CK3	199,035	62,467,550	314	519			
	Average	180,837	60,006,622	334	588			
	G48h1	184,369	63,031,414	342	600			
	G48h2	204,199	66,200,481	324	544			
	G48h3	178,076	59,697,081	335	599			
	Average	188,881	62,976,325	334	581			
	G96h1	206,643	67,437,004	326	537			
	G96h2	208,832	66,755,999	320	525			
	G96h3	208,483	66,220,813	322	536			
	Average	207,986	67,137,939	323	533			
Unigene	CK1	95,594	59,703,496	625	1,260	95,594	26,277	69,317
	CK2	127,820	76,152,752	596	1,171	127,820	35,504	92,316
	CK3	118,263	67,850,541	574	1,154	118,263	38,092	80,171
	Average	113,892	67,902,263	598	1,195	113,892	33,291	80,601
	G48h1	120,623	72,427,471	600	1,175	120,623	33,584	87,039
	G48h2	127,545	75,575,056	593	1,180	127,545	38,108	89,437
	G48h3	112,385	69,235,942	616	1,237	112,385	32,742	79,643
	Average	120,184	72,412,823	603	1,197	120,184	34,811	85,373
	G96h1	132,647	79,390,962	599	1,182	132,647	37,879	94,768
	G96h2	131,163	79,154,740	603	1,210	131,163	39,713	91,450
	G96h3	132,446	82,054,570	620	1,255	132,446	40,649	91,797
	Average	132,085	80,200,091	607	1,216	132,085	39,414	92,672
	All	203,443	182,572,023	897	1,523	203,443	89,153	114,290

Note: The total consensus sequences are all unigenes obtained by the sequence assembly; distinct clusters indicate different unigene clusters, and unigenes clustered in one group shared more than 70% similarity with each other, indicating that they might have been derived from the same gene ancestor or are homologous genes; distinct singletons indicate that the unigenes were derived from a single gene.

Table 3. Statistics related to the unigene BLAST annotation against six different databases.

Database	Number of Annotated Unigenes	Annotated Unigene Ratio (%)
NR	81,322	94.94%
NT	50,836	59.35%
Swiss-Prot	53,382	62.32%
KEGG	50,844	59.36%
COG	31,315	36.56%
GO	55,986	65.36%

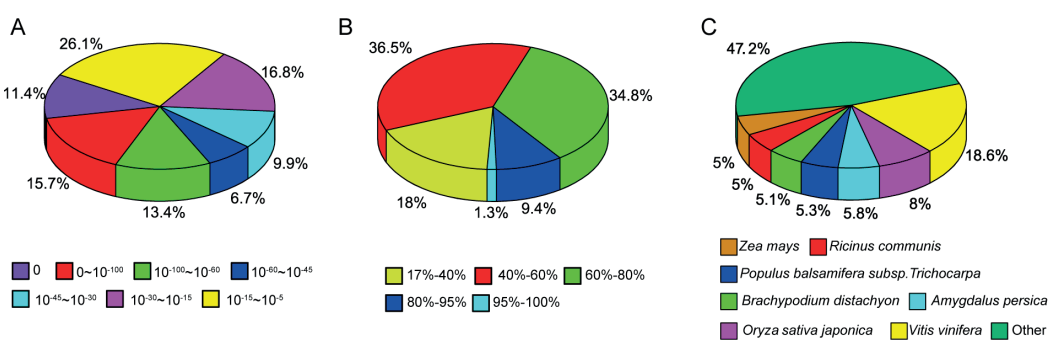

Figure 2. Similarity profile of the unigenes versus the NR database search. (A) E-value distribution of the BLAST hits of each unigene with a cutoff E-value of 1.0 e^{-5}. (B) Similarity distribution of the top BLAST hit of each unigene. (C) Species distribution of the top BLAST hit of each unigene in the NR database.

exploring the biological process category, most unigenes were involved in "cellular process" (36,178 unigenes), "metabolic process" (35,686 unigenes), and "single-organism process" (22,783 unigenes). Additionally, in the cellular component category, many unigenes were assigned to "cell" (41,522 unigenes), "cell part" (41,466 unigenes), "membrane" (20,967 unigenes), and "organelle" (33,221 unigenes). Regarding the molecular function category, most corresponding unigenes were assigned to "binding" (25,110 unigenes) and "catalytic activity" (28,334 unigenes) (Figure 3A).

In addition to the GO analysis, a COG analysis was used to further define the unigene functions (Tatusov et al., 2000). Of the 81,322 NR hits,

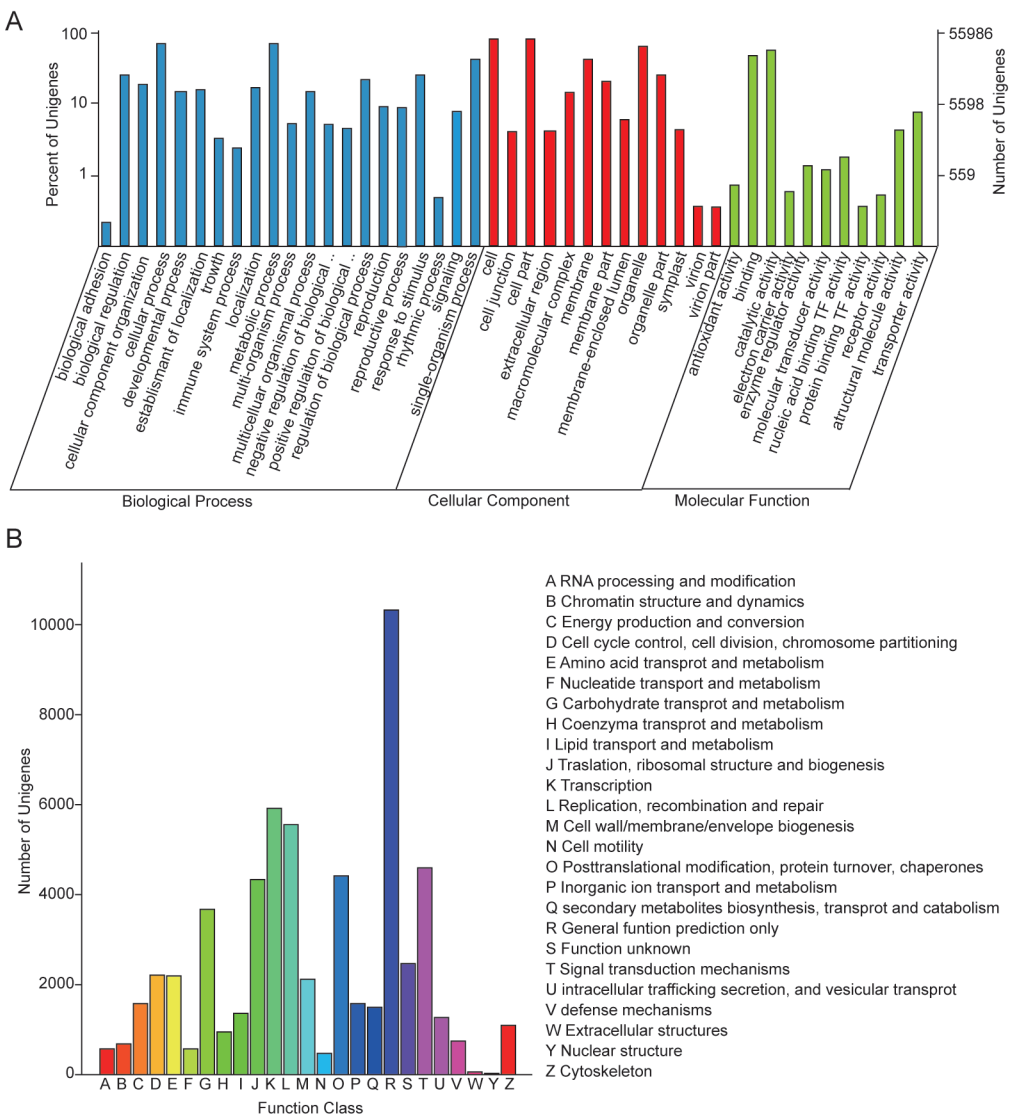

Figure 3. Functional classification of the assembled unigenes. (A) Functional classification of the assembled unigenes based on Gene Ontology (GO) categorization. Unigenes were assigned to the following three categories: biological processes, cellular components, and molecular function. The X-axis indicates the subcategories; the left Y-axis indicates the percentage of unigenes related to the total GO terms present; the right Y-axis indicates the numbers of unigenes related to the total GO terms present. (B) Histogram of the Clusters of Orthologous Groups (COG) classification. All unigenes were aligned with the COG database to predict and classify their possible functions.

31,315 hits were aligned with 25 COG classifications. Across all COG categories, the cluster "general function prediction only" (10,371; 17.4013%) represented the largest group, followed by "transcription" (5,892; 0.0989%) and "replication, recombination and repair" (5,529; 0.0928%). Only small proportions of the unigenes were classified as "extracellular structures" (33; 0.0006%) and "nuclear structure" (15; 0.0003%) (Figure 3B).

Analysis of differentially expressed genes (DEGs)

To identify genes that display significant expression dynamics in drought plants, the DEGs among the samples were selected based on the fold-change values with a log2 (fold change) >1 or a log2 (fold change) <–1 and statistical significance (P<0.05) using the R package edger (Robinson et al., 2010). We identified 253/266 up/downregulated DEGs between the control (CK) and drought-treated for 48 hours (G48h) samples (Figure 4A), 263/11013 up/downregulated DEGs between the CK and drought-treated for 96 hours (G48h) samples (Figure 4B), and 69/7 up/downregulated DEGs between the two drought-treated samples (Figure 4C).

Functional categorization of stress-regulated DEGs

To classify the function of the differentially expressed genes under drought stress, GO assignments were conducted. Generally, the most enriched GO results of the DEGs were categorized into the following three main categories: biological process (BP), cellular component (CC), and molecular function (MF); the distributions of the gene numbers in each category are shown.

The DEGs between CK and G48h were enriched in the BP of lipid biosynthetic process, photosynthesis, isoprenoid biosynthetic and metabolic processes, isopentenyl diphosphate biosynthetic and metabolic processes, cellular aldehyde, and glyceraldehyde-3-phosphate; MF of oxidoreductase and oxalate decarboxylase activity, NADPH dehydrogenase activity, RNA binding, monooxygenase activity, GDP-mannose 3,5-epimerase activity, and chlorophyll

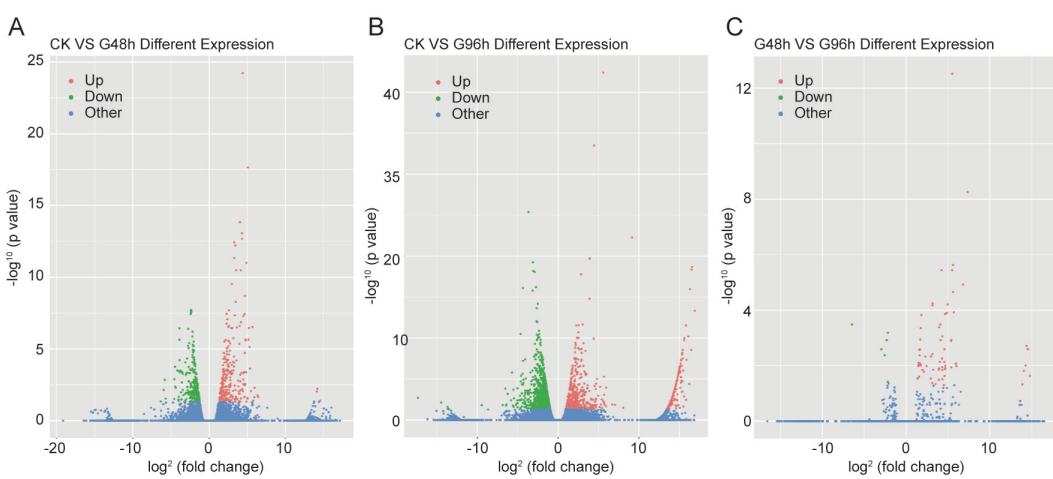

Figure 4. Differential gene expression analysis under different drought stress conditions in *A. mongolicus* root samples. Mean plot of the DEG analysis between CK and G48h (A); CK and G96h (B); and G48h and G96h (C). The red dots show significantly upregulated genes, the green dots show significantly downregulated genes, and the blue dots indicate no significant differences.

binding activity; and CC of plastid, chloroplast, and thylakoid (Figure 5A). In contrast to the G48h samples, in the G96h samples, more DEGs were involved in the BP of metabolic process; MF of symporter, sugar, and carbohydrate transmembrane transporter activity; and CC of the vesicle, plasma membrane, and extracellular region (Figure 5B). These results imply that with the prolongation of the drought treatment time, more plant protection mechanisms are activated. Furthermore, the DEGs between the two treated samples associated with biosynthetic processes and various metabolic processes were screened (Figure 5C). These results indicate that energy-related metabolic processes and secondary metabolism were active in *A. mongolicum Regel* roots under drought stress.

To further understand the biological pathways active in the roots of PEG-treated *A. mongolicum Regel*, all DEGs were mapped to the reference canonical pathways in KEGG. The same results as above were obtained in the CK versus G48h samples. The most enriched DEGs were accumulated during secondary biosynthesis, metabolism-related pathways, carbohydrate biosynthesis, metabolism-related pathways, and photosynthesis-related pathways. However, in the G96h samples, cyan amino, valine, leucine, and isoleucine were not collected during carbohydrate biosynthesis or in metabolism-related pathways (Figure 6A). Regarding the G48h versus G96h samples, glutathione and alpha-linolenic acid metabolism represented the most dominant pathways (Figures 6B). Surprisingly, this was not evident in the comparisons between CK and G48h and CK and G96h (Figures 6A–6C). These results indicate that following drought treatment in *A. mongolicum Regel* roots, energy metabolism and secondary metabolism are involved in the early stages; however, over time, an increasing number of biological processes are mobilized.

Validation of differentially regulated transcripts

To prove the results of the high-throughput sequencing, five transcripts, which coded four functional proteins and one transcription factor, were selected for experimental validation and expression profiling. It has been reported that these five genes (chlorophyll a-b binding protein, *SWEET14*, LTP, LEA protein, and MYB) contribute to drought resistance in plants (Seki et al., 2010; Artur et al., 2019; Jadid et al., 2018). All analyzed genes were found to be downregulated after 48 hours of the drought treatment. However, the expression of the MYB transcription factor and LTP increased 96 hours after PEG stress (Figure 7).

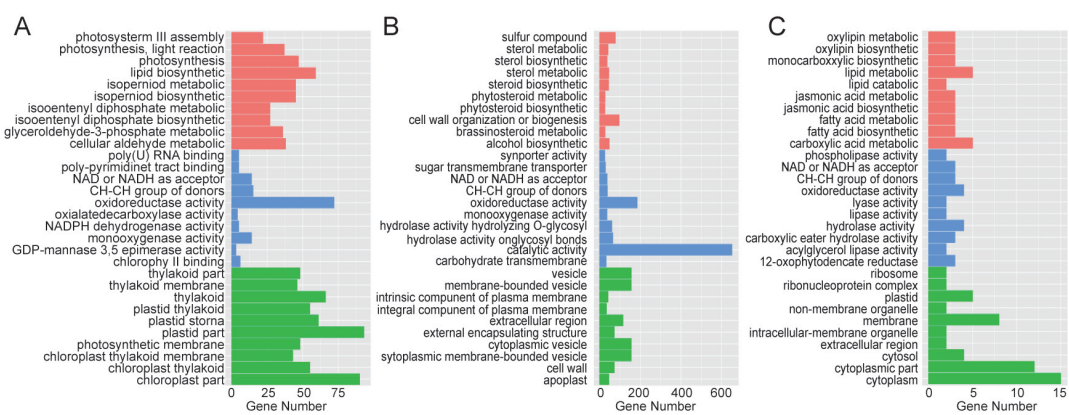

Figure 5. Functional classification of differentially expressed unigenes (DEGs) based on Gene Ontology (GO) categorization. GO terms enriched in the differential expression analysis between (A) CK and G48h; (B) CK and G96h; and (C) G48h and G96h. Red: biological process; green: cellular component; blue: molecular function.

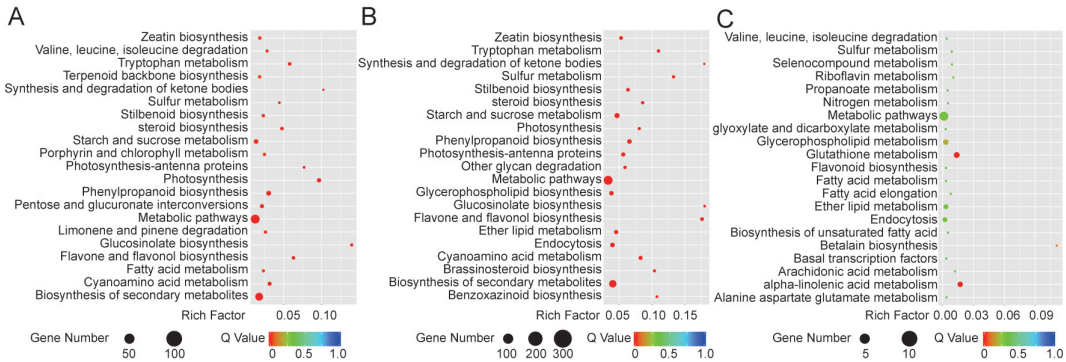

Figure 6. Functions of DEGs classified by KEGG pathways. Scatterplot of the KEGG pathway enrichment statistics. (A) CK versus G48h; (B) CK versus G96h; and (C) G48h versus G96h. The rich factor is the ratio of the differentially expressed gene numbers annotated in this pathway term to all gene numbers annotated in this pathway term. A greater rich factor indicates greater intensiveness. The Q-value is a corrected *P*-value ranging from 0~1; a lower value indicates greater intensiveness. The top 20 pathway terms enriched in the KEGG database are displayed.

Discussion

Global warming, continuous population increases, and decreases in arable land have led us to seek additional methods to improve our living conditions. To adapt to water-scarce environments, desert plants have developed multiscale, stress-tolerant strategies to mitigate the effects of stress injuries (Bhardwaj et al., 2015; Banerjee & Roychoudhury, 2016). Therefore, analyzing the adaptation mechanisms of desert plants and uncovering new drought-resistance genes are among the many possible ways to solve this problem. Although a transcriptomic analysis of Allium has been previously performed (Tsukazaki et al., 2015; 2014), the molecular mechanisms underlying drought resistance in *A. mongolicum Regel* are still unclear. Here, using de novo RNA-seq assembly, a large amount of sequencing data was produced in *A. mongolicum Regel*, resulting in significant coverage of the transcriptome in roots.

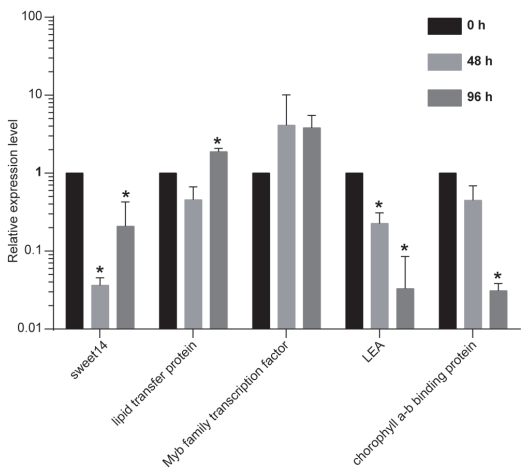

Figure 7. Relative expression level analysis of selected transcripts as determined by qPCR. Expression profiling of differentially expressed transcripts was analyzed using quantitative real-time PCR. The relative expression level (Y-axis) was calculated using the ΔΔCt method. *A. mongolicum Regel* seedlings were subjected to drought stress for 48 hours and 96 hours. The mean of three independent biological replicates is presented. The Allium actin gene was used as an internal control; the values are the means ± standard errors (SEs) of three independent biological replicates (n=3; **P*<0.05; ***P*<0.01; *t*-test).

Many complex cellular, biochemical, and physiological processes participate in plant responses to drought stress (Raney et al., 2014), such as osmotic homeostasis or regulation, stress repair and defense, detoxification, and growth inhibition (Yuan et al., 2016). Moreover, during the drought-response process, an assortment of genes with diverse functions is induced, and most of their products may be directly or indirectly involved in the stress response and drought tolerance (Ma et al., 2017; Zhang et al., 2017). In this work, many differentially expressed genes were revealed in drought-stressed *A. mongolicum*

Regel as identified through pairwise comparisons of the gene expression levels in the CK, G48h, and G96h samples (Figure 4). Following the functional annotation, the differential gene expression in response to drought was mostly involved in metabolic pathways, the biosynthesis of secondary metabolites and starch and sucrose metabolism (Figures 6B and 6C), suggesting that genes participating in energetic and secondary metabolism were heavily influenced by drought stress in *A. mongolicum Regel*. Among the DEGs anticipated by RNA-seq, five genes were validated by qPCR. Overall, our conclusions were consistent with those in previous reports.

During the regulation of drought stress, these signaling pathways play an important role. Sugar transporters are essential for osmotic pressure balance under stress conditions. Additionally, lipid-transfer proteins and fatty acid metabolism-related genes play important roles in the repair of stress-induced membrane damage and membrane lipid composition changes (Seki et al., 2010). Secondary metabolite-mediated metabolic processes represent another pathway that can help plants overcome drought conditions. In this work, genes involved in brassinosteroid metabolic processes, steroid metabolic processes, and jasmonic acid metabolic processes were found to be highly enriched in drought plants. Meanwhile, the KEGG pathway analysis results also suggested that metabolic pathways and the biosynthesis of secondary metabolites were significantly enriched in response to drought. These results imply that ABA-mediated signaling pathways and secondary metabolism may help desert plants maintain water balance and membrane integrity under water deprivation.

Furthermore, to prove the authenticity of the assembly results, five differentially expressed transcripts were selected and validated using qPCR, and the results were in agreement with the RNA sequencing results. It has been reported that sugar transporters adjust osmotic pressure under stress conditions by transporting sugars through the plasma membrane and tonoplast (Zandalinas et al., 2016). The bidirectional sugar transporter SWEET14, which is a sugar transporter highly expressed in drought plants, is considered associated with effective root carbohydrate accumulation and the root elongation/enlargement performance of plants under drought stress (Yıldırım et al., 2018). In addition to sugar transporters, transcription factors (TFs), such as DREB, bZIP, zinc-finger, MYB, and the NAC protein, are involved in stress conditions (Zhao et al., 2018). Based on the results of this study, many TFs that respond to drought were identified, particularly those involved in secondary metabolism, transport, photosynthesis, and lipid metabolism. For example, the transcription factor MYB is an important regulator underlying plant responses to environmental stress (Zhao et al., 2018). The increased expression of the *MYB* gene in the PEG-treated roots suggests that *MYB* genes might contribute to drought resistance in *A. mongolicum Regel*. Similar to the *MYB* gene, the functional gene *LIPID TRANSFER PROTEIN* (*LTP*), which is tightly associated with enhanced plant drought resistance, also demonstrated a similar gene expression profile consistent with that involved in the process of drought stress (Jadid N et al., 2018). Meanwhile, the expression trends of the *LTP* gene echoed the KEGG analysis results, which showed that genes involved in the "lipid metabolic and biosynthetic process" were differentially expressed between the 48-hour drought seedlings and 96-hour drought seedlings. Finally, the *LEA* gene, which is known to be expressed in response to water loss in vegetative tissues of desiccation-tolerant species, was decreased in the 96-hour drought seedlings (Artur et al., 2018).

In summary, this annotation provides a valuable resource for investigating specific processes, functions, and pathways involved in drought tolerance in *A. mongolicum Regel* and further facilitates the identification of novel drought-resistance genes in *A. mongolicum Regel*.

Conclusion

Recent studies have suggested that numerous genes are differentially expressed in plants in response to drought stress. Thus, dissecting the root transcriptome may provide novel insight into the response of desert plants to drought conditions. In this study, a comprehensive root transcriptome of *A. mongolicum Regel* was constructed using de novo RNA-seq assembly. Using pairwise comparisons of drought plants and nondrought plants, the differentially expressed transcripts were mostly involved in secondary metabolism, transport, photosynthesis, and lipid metabolism. Following further analysis of the data, we found that several genes that participate in the ABA-mediated signaling pathway and secondary metabolism were upregulated in the roots of *A. mongolicum Regel*, suggesting that these genes play important roles under drought stress. Our data further showed that different genes and expression levels were induced by different treatment times under drought stress. These results may provide clues that could help us better understand the drought resistance mechanisms of *A. mongolicum Regel* and further aid in the selection of drought-resistant genes, which might be utilized to develop drought-tolerant plants.

Competing interests

The authors declare that they have no competing interests.

Availability of supporting data

The data discussed in this publication were deposited in NCBI and are accessible through series accession number PRJNA526320.

Resumen

N. Tang, B.-J. Zhang, P. Yang, L.-K. Wang, y Q.-Z. Chen. 2022. Transcriptoma radicular de *Allium mongolicum* Regel en condiciones de estrés por sequía. Int. J. Agric. Nat. Resour. 22-35. *Allium mongolicum* Regel es un tipo de planta desértica que muestra una gran resistencia a la erosión eólica, la sequía y las bajas temperaturas; se distribuye ampliamente en las tierras desérticas de China. Comprender los mecanismos moleculares subyacentes a la resistencia a la sequía de *A. mongolicum* Regel es esencial para descubrir genes resistentes a la sequía y mejorar este rasgo beneficioso en otras plantas. Aquí, el análisis de ensamblaje de ARN-Seq de novo se llevó a cabo utilizando las raíces de plántulas de 1 mes de edad en respuesta al estrés por sequía. Utilizando comparaciones por pares de plantas no tratadas (CK), plantas de 48 horas (G48h) y plantas de 96 horas (G96h), se obtuvieron un total de 2.211 unigenes (DEGs) expresados diferencialmente. Mientras tanto, usando la anotación funcional, estos DEGs fueron implicados principalmente en la membrana de plasma, la fotosíntesis, y el metabolismo de carbohidrato secundario. Por otra parte, los genes implicados en el camino ABA-mediado de la señalización y en metabolismo secundario para arriba fueron regulados para arriba en raíces. Estos resultados sugieren que las plantas del desierto pueden utilizar la señalización y las vías metabólicas secundarias como respuestas adaptativas al estrés por sequía. Colectivamente, este trabajo ayudará a comprender los mecanismos moleculares subyacentes a la capacidad de *A. mongolicum* Regel para responder al estrés por sequía y ayudar en la selección de nuevos genes de tolerancia a la sequía.

Palabras clave: *Allium mongolicum*, estrés por sequía, mecanismo, molecular, raíz, transcriptoma.

References

Artur, M., Zhao, T., Ligterink, W., Schranz, M.E., & Hilhorst, H.W.M. (2018). Dissecting the genomic diversification of Late Embryogenesis Abundant (LEA) protein gene families in plants. *Genome Biology and Evolution*.

Artur, M.A.S., Zhao, T., Ligterink, W., Schranz, E., & Hilhorst, H.W.M. (2019). Dissecting the Genomic Diversification of Late Embryogenesis Abundant (LEA) Protein Gene Families in Plants. *Genome Biol Evol*, 11(2), 459–471. doi:10.1093/gbe/evy248.

Banerjee, A., & Roychoudhury, A. (2016). Group II late embryogenesis abundant (LEA) proteins: structural and functional aspects in plant abiotic stress. *Plant Growth Regulation*, 79(1), 1–17.

Bhardwaj, A.R., Joshi, G., Kukreja, B., Malik, V., Arora, P., Pandey, R., et al. (2015). Global insights into high temperature and drought stress regulated genes by RNA-Seq in economically important oilseed crop Brassica juncea. *Bmc Plant Biology*, 15(1), 9.

Conesa, A., Gotz, S., Garciagomez, J.M., Terol, J., Talon, M., & Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21(18), 3674–3676.

De Hoon, M.J.L., Imoto, S., Nolan, J., & Miyano, S. (2004). Open source clustering software. *Bioinformatics*, 20(9), 1453–1454.

Grabherr, M., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., et al. (2011a). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29(7), 644–652.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., et al. (2011b). Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology*, 29(7), 644–652.

Hu, J., Ma, Q., Kumar, T., Duan, H.-R., Zhang, J.-L., Yuan, H.-J., et al. (2016). ZxSKOR is important for salinity and drought tolerance of Zygophyllum xanthoxylum by maintaining K+ homeostasis. *Plant growth regulation*, 80(2), 195–205.

Iseli, C., Jongeneel, C.V., & Bucher, P. (1999) 'ESTScan: A Program for Detecting, Evaluating, and Reconstructing Potential Coding Regions in EST Sequences' *intelligent systems in molecular biology*. pp. 138–148.

Jadid, N., Estiasih, E., Saputro, T.B., Purwani, K.I., Hidayati, D., Permatasari, E.V., et al. (2018). Expression pattern of drought-responsive genes in burley tobacco under in vitro water deficit. *Journal of Physics: Conference Series* 1040(1), 012004.

Khan, M.R., Khan, I., Ibrar, Z., Leon, J., & Naz, A.A. (2017). Drought-responsive genes expressed predominantly in root tissues are enriched with homotypic cis -regulatory clusters in promoters of major cereal crops. *Crop Journal*, 5(3), 195–206.

Li, W., Nguyen, K.H., Chu, H.D., Van Ha, C., Watanabe, Y., Osakabe, Y., et al. (2017a). The karrikin receptor KAI2 promotes drought resistance in Arabidopsis thaliana. *PLOS Genetics*, 13(11).

Li, X., Han, H., Chen, M., Yang, W., Liu, L., Li, N., et al. (2017b). Overexpression of OsDT11 , which encodes a novel cysteine-rich peptide, enhances drought tolerance and increases ABA concentration in rice. *Plant Molecular Biology*, 93, 21–34.

Muqier, Qi, S., Wang, T., Chen, R., Wang, C., & Ao, C. (2017). Effects of flavonoids from Allium mongolicum Regel on growth performance and growth-related hormones in meat sheep. *Anim Nutr*, 3(1), 33–38. doi:10.1016/j.aninu.2017.01.003.

Pointing, S.B., & Belnap, J. (2012). Microbial colonization and controls in dryland systems. *Nature Reviews Microbiology*, 10(8), 551–562.

Qin, T., Zhao, H., Cui, P., Albesher, N., & Xiong, L. (2017). A Nucleus-localized Long Non-Coding RNA Enhances Drought and Salt Stress Tolerance. *Plant Physiology*, 175(3), 1321–1336.

Raney, J.A., Reynolds, D.J., Elzinga, D.B., Page, J.T., Udall, J.A., Jellen, E.N., et al. (2014). Transcriptome Analysis of Drought Induced Stress in Chenopodium Quinoa. *American Journal of Plant Sciences*, 05(03), 338–357.

Rodriguezcaballero, E., Belnap, J., Budel, B., Crutzen, P.J., Andreae, M.O., Poschl, U., et al. (2018). Dryland photoautotrophic soil surface communi-

ties endangered by global change. *Nature Geoscience*, 11(3), 185–189.

Saldanha, A.J. (2004). Java Treeview---extensible visualization of microarray data. *Bioinformatics*, 20(17), 3246–3248.

Shao, Y., Zhang, Y., Wu, X., Bourque, C.P.A., Zhang, J., Qin, S., et al. (2016). Relating historical vegetation cover to aridity index patterns in the greater desert region of northern China: Implications to planned and existing restoration projects. *Biogeosciences Discussions*, 1–22.

Shinozaki, K., & Yamaguchi-Shinozaki, K. (2006). Gene networks involved in drought stress response and tolerance. *Journal of Experimental Botany*, 58(2), 221–227.

Sultana, R., Khurram, B., Akihiro, M., Maho, T., & Motoaki, S. (2016). Transcriptomic Analysis of Soil-GrownArabidopsis thalianaRoots and Shoots in Response to a Drought Stress. *Frontiers in Plant Science*, 7(180), 180.

Tatusov, R.L., Galperin, M.Y., Natale, D.A., & Koonin, E.V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res*, 28(1), 33–6. http://www.ncbi.nlm.nih.gov/pubmed/10592175.

Tsukazaki, H., Yaguchi, S., Sato, S., Hirakawa, H., Katayose, Y., Kanamori, H., et al. (2015). Development of transcriptome shotgun assembly-derived markers in bunching onion (Allium fistulosum). *Molecular Breeding*, 35(1), 55.

Wang, G., Hui, Z., Li, F., Zhao, M., Zhang, J., & Wang, W. (2010). Improvement of heat and drought photosynthetic tolerance in wheat by overaccumulation of glycinebetaine. *Plant Biotechnology Reports*, 4(3), 213–222.

Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., et al. (2006). WEGO: a web tool for plotting GO annotations. *Nucleic Acids Research*, 34, 293–297.

Yıldırım, K., Yağcı, A., Sucu, S., & Tunç, S. (2018). Responses of grapevine rootstocks to drought through altered root system architecture and root transcriptomic regulations. *Plant Physiology & Biochemistry*, 127, 256–268.

Yuan, L.L., Zhang, M., Yan, X., Bian, Y.W., Zhen, S.M., & Yan, Y.M. (2016). Dynamic Phosphoproteome Analysis of Seedling Leaves in Brachypodium distachyon L. Reveals Central Phosphorylated Proteins Involved in the Drought Stress Response. *Scientific Reports*, 6, 35280.

Zandalinas, S.I., Rivero, R.M., Martínez, V., Gómez-Cadenas, A., & Arbona, V. (2016). Tolerance of citrus plants to the combination of high temperatures and drought is associated to the increase in transpiration modulated by a reduction in abscisic acid levels. *Bmc Plant Biology*, 16(1), 105.

Zhang, C., Preece, C., Filella, I., Farré-Armengol, G., & Peñuelas, J. (2017). Assessment of the Response of Photosynthetic Activity of Mediterranean Evergreen Oaks to Enhanced Drought Stress and Recovery by Using PRI and R690/R630. *Forests*, 8(10), 386. http://www.mdpi.com/1999-4907/8/10/386.

Zhao, Y., Cheng, X., Liu, X., Wu, H., Bi, H., & Xu, H. (2018). The wheat MYB transcription factor TaMYB31 is involved in drought stress responses in Arabidopsis. *Frontiers in plant science*, 9.

Zhou, Y., Gao, F., Liu, R., Feng, J., & Li, H. (2012). De novo sequencing and analysis of root transcriptome using 454 pyrosequencing to discover putative genes associated with drought tolerance in Ammopiptanthus mongolicus. *BMC Genomics*, 13(1), 266–266.